

Verwant met verband: Ruimte, Recht en Wetenschap

Vriendenbundel voor prof. mr. J. Struiksma

Onder redactie van:
mr. dr. P.J. (Pim) Huisman
prof. dr. A.R. (Richard) Neerhof
prof. mr. F.J. (Frank) van Ommeren

Eerste druk



's-Gravenhage - 2019

Advies van een deskundige of algoritme?

De toetsing van ‘black box’-besluiten door de bestuursrechter

Mr. N.H. (Heleen) van Amerongen & prof. mr. Y.E. (Ymre) Schuurmans

1 Inleiding

Burgers en bedrijven krijgen regelmatig te maken met besluiten die bestuursorganen met ‘hulpmiddelen’ hebben voorbereid. Deze hulpmiddelen kunnen adviezen van één of meerdere deskundigen zijn, maar ook bepaalde (digitale) modellen, programma’s of software. De inzet van deze hulpmiddelen stelt bestuursorganen in staat om inzicht te krijgen in specialistische, wetenschappelijke of technisch complexe materie, die zij anders niet of maar ten dele zouden kunnen ontwikkelen. Sterker nog, zij kunnen zelfs volledig of ten dele van deze hulpmiddelen afhankelijk zijn om op grote schaal besluiten op basis van *state-of-the-art* kennis te nemen. De keerzijde is dat dergelijke complexe besluiten ondoorzichtig en oncontroleerbaar kunnen worden, omdat onduidelijk is hoe het besluit tot stand is gekomen en waarop het besluit inhoudelijk is gebaseerd. Ze kunnen als *black boxes* worden ervaren.¹

De bestuursrechter die effectieve rechtsbescherming wil bieden, zoekt naar manieren om die *black boxes* te openen. Dit past binnen een meer algemene trend van de bestuursrechter om bestuurlijke besluitvorming indringender te toetsen, ook als bestuursorganen beslissingsruimte hebben.² Belangrijke drijfveren achter deze ontwikkeling zijn de (rechts)bescherming van de burger, maar ook de behoefte aan versterkte controle op bestuurshandelen, nu de democratische legitimatie van besluiten vaak flinterdun is.³

Die behoefte om de black box te openen doet zich in de digitale wereld sterk gevoelen. Een besluit dat geautomatiseerd tot stand komt, wordt voor een groot deel bepaald door

1 Deze term wordt vooral gebruikt in de context van algoritme-gedreven besluitvorming, zie F. Pasquale, *The Black Box Society. The Secret Algorithms That Control Money and Information*, Cambridge: Harvard University Press 2015.

2 R. Ortley & W.S. Zorg, ‘Van marginale rechterlijke toetsing naar toetsing op maat: einde van een geconditioneerde respons?’, *AA* 2018, afl. 1, p. 20-25; R. Kegge, ‘Indringende(re) toetsing door de bestuursrechter’, *JBplus* 2018, afl. 4.

3 Zie o.a. E.M.H. Hirsch Ballin, ‘Dynamiek in de bestuursrechtspraak’, in: E.M.H. Hirsch Ballin, R. Ortley & A. Tollenaar, *Rechtsontwikkeling door de bestuursrechter* (preadviezen VAR, VAR-reeks 154), Den Haag: Bju 2015, p. 7-58 en de Conclusie van A-G Widdershoven van 22 december 2017, ECLI:NL:RVS:2017:3557.

een algoritme dat op basis van daarin vastgelegde instructies inputdata omzet in een beslissing. De rechter wil dan weten welke data zijn gebruikt, van welke veronderstellingen het programma uitgaat en welke beleidsmatige keuzes zijn gemaakt. De breed uitgemeten PAS-uitspraak biedt daar een duidelijk voorbeeld van.⁴ Maar is het openen van de black box van (deels) geautomatiseerde besluitvorming van fundamenteel andere orde dan het openen van de black box van deskundigenadvisering? Hebben we andere normerende beginselen nodig of kunnen we voor een groot deel varen op de wijze waarop we in het algemeen met deskundigenadvisering omgaan?⁵ Daarbij speelt de vraag of de digitale wereld werkelijk van de analoge verschilt.

Wij stellen de volgende onderzoeksvraag centraal: *Aan de hand van welke normen toetst de bestuursrechter het deskundigenadvies dan wel het beslissysteem en kunnen overeenkomsten dan wel verschillen worden verklaard vanuit waarden als rationaliteit en 'accountability'?* Onder rationaliteit verstaan we de mate waarin de deskundige dan wel het beslissysteem bijdraagt aan een accurate feitenvaststelling en -waardering en zodoende materiële waarheidsvinding bevordert. 'Accountability' is een vooral in het internationale debat gebruikt concept, dat eisen stelt aan deskundigenadvisering en algoritme-gedreven besluitvorming.⁶ Met 'accountability' doelen we op de mate waarin het bestuursorgaan verantwoordelijkheid kan nemen voor de beslissing. De inzichtelijkheid en controleerbaarheid van de beslissing speelt daarbij een grote rol. Beide waarden kunnen gediend zijn met transparantie.⁷ Transparantie kan rationaliteit versterken door bijvoorbeeld een zinnig bewijsdebat tussen partijen te bevorderen en accountability door beslissingen meer inzichtelijk en controleerbaar te maken.

In deze bijdrage beschrijven we enkele casus op verschillende deelterreinen die we indelen op een glijdende schaal. Paragraaf 2 beschrijft de bestuursrechtelijke eisen die aan algoritme-gedreven besluitvorming worden gesteld. Paragraaf 3 behandelt de tussensituatie waarin de deskundige mens adviseert, maar waarbij zijn oordeel mede wordt gevormd door de computersystemen waarin hij de beoordeling moet invoeren. Paragraaf 4 beziet de bestuursrechtelijke normering van de advisering door de 'traditionele' deskundige. In paragraaf 5 duiden we overeenkomsten en verschillen in de normering en waarden die in het licht van rationaliteit en 'accountability'. We sluiten af met een conclusie.

De bijdrage is verkennend van aard, waarin we de gelijksoortige dan wel uiteenlopende benaderingswijze van de bestuursrechter identificeren en in een breder kader plaatsen.

4 ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, *M&R* 2017/84, m.nt. M.M. Kaajan, *Computerrecht* 2017/256, m.nt. B.M.A. van Eck, *Gst.* 2017/170, m.nt. B. Assink.

5 *Stcrt.* 2018, 50 999 (Advies Raad van State betreffende digitalisering), p. 12-14. Zie ook M. Oswald, 'Algorithm-assisted-decision-making in the public sector: framing the issues using administrative law rules governing discretionary power', *Phil. Trans. R. Soc. A* 2018, 376:20170359 en C. Coglianese & D. Lehr, 'Regulating by robot: Administrative Decision Making in the Machine-Learning Era', *The Georgetown Law Journal* 2017, p. 1147-1223.

6 Zie bijvoorbeeld N. Diakopoulos, 'Algorithmic Accountability', *Digital Journalism* 2015 (3), nr. 3, p. 402 en M. Ambrus, K. Arts, E. Hey & H. Raulus, 'The role of experts in International and European decision-making processes: setting the scene', in: M. Ambrus e.a. (red.), *The role of 'Experts' in International and European Decision-Making Processes*, Cambridge: University Press 2014, p. 7-8.

7 A.W.G.J. Buijze, 'Het transparantiebeginsel naar Nederlands recht: een visie geïnspireerd op het EU-recht', *JBplus* 2016/18, afl. 4.

De casus hebben we geselecteerd op de grote mate van invloed van het algoritme of de deskundige op het besluit, waarbij de bestuursrechter richtinggevende uitspraken heeft gedaan over de eisen die het bestuursrecht stelt. De literatuuranalyse is eveneens verkennend van aard, waarbij het doel is de bestuursrechtelijke normering niet alleen in het licht van de Awb te plaatsen, maar ook in bredere concepten van controle van het openbaar bestuur, waarin de ratio van de maatstaven tot uitdrukking komt. Wij menen dat deze opzet goed past in de onderzoekslijn en -traditie van Jan Struiksma: het beziet op licht eigennuttige wijze een ontwikkeling in maatschappij en recht, niet alleen vanuit de Awb of het bijzondere bestuursrecht, maar gaat op zoek naar de ratio van die normen (en *common sense*) en beziet of de normen aansluiten bij de probleemanalyse.⁸ Bovendien ligt de bijdrage in het hart van Jans belangstelling voor digitalisering, rechtsbescherming en bewijsrecht.⁹

2 Het deskundige systeem

2.1 (Deels) geautomatiseerde besluitvorming

Geheel of gedeeltelijk geautomatiseerde besluitvorming wordt door computersystemen uitgevoerd op basis van algoritmes en vaak grote hoeveelheden data. Hoe toetst de bestuursrechter dit type besluitvorming? Om dit in kaart te brengen staan we kort stil bij het Programma Aanpak Stikstof (PAS, 2.2) en de WOZ-beschikking (2.3). Vooral de (deels) geautomatiseerde besluitvorming binnen het PAS kreeg brede aandacht, omdat de Afdeling in die zaak voor het eerst expliciet een toetsingskader ontwikkelde.¹⁰ Dit toetsingskader bleek bestuursrechtbreed toepassing te krijgen: ook de Hoge Raad heeft dit inmiddels toegepast in het kader van de WOZ-beschikking.

2.2 Het Programma Aanpak Stikstof en AERIUS

Nederland staat voor de enorme opgave om bepaalde stikstofgevoelige natuurgebieden te beschermen tegen de effecten van stikstofoverbelasting en tegelijkertijd economische activiteiten de ruimte te geven. Voor deze afweging is een integrale aanpak bedacht in het PAS, dat sinds 2015 geldt.¹¹ Dit programma bestaat onder andere uit een meldings- en vergunningsverleningssysteem voor activiteiten die stikstofdepositie in beschermde gebieden veroorzaken. Op grond van het voorzorgsbeginsel van de Habitatrichtlijn is het nodig per plan of project een passende beoordeling te maken die moet *verzekeren*

8 Zie bijv. J. Struiksma, 'Het wetsvoorstel Omgevingswet: de beperkingen van een wetenschappelijke beoordeling', *TBR* 2014/142.

9 Zie o.a. J. Struiksma, 'Digitaal stelsel Omgevingswet: uithuilen en opnieuw beginnen', *WPNR* 2018, p. 58-64; J. Struiksma, *Het geschil als spil. De empirische cyclus van de rechtswetenschap*, Zaandam: Struiksma 2012 en zijn grote betrokkenheid als initiator en promotor van het promotieonderzoek Y.E. Schuurmans, *Bewijslastverdeling in het bestuursrecht* (diss. Amsterdam, VU), Deventer: Kluwer 2005.

10 ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, *M&R* 2017/84, m.nt. M.M. Kaajan, *Computerrecht* 2017/256, m.nt. B.M.A. van Eck, *Gst.* 2017/170, m.nt. B. Assink. Naast deze uitspraak was er een daaraan gerelateerde uitspraak die dag (ECLI:NL:RVS:2017:1260).

11 Zie voor een uitgebreide uitleg over dit programma: R.H.W. Frins, 'PAS op de plaats... en nu? (Deel I)', *TBR* 2017/93, p. 586-596.

dat de kwaliteit van stikstofgevoelige natuurgebieden niet verslechtert.¹² Deze passende beoordeling moet gebaseerd zijn op de beste wetenschappelijke kennis en mag geen redelijke wetenschappelijke twijfel laten bestaan over de vraag of het project de natuurlijke kenmerken van het natuurgebied zal aantasten.¹³ De ministers van Economische zaken en Infrastructuur en Milieu¹⁴ hebben aan deze passende beoordeling een grote hoeveelheid onderzoeken ten grondslag gelegd en zijn in het PAS van verschillende aannames en marges uitgegaan om onzekerheden gedurende de looptijd van het programma van zes jaar op te vangen. Het Hof van Justitie van de EU uitte in antwoord op prejudiciële vragen van de Afdeling zorgen hierover en heeft de nationale rechter opgedragen de wetenschappelijke deugdelijkheid van het PAS nog eens grondig en volledig te toetsen.¹⁵

De uitkomsten van de passende beoordeling zijn verwerkt in de rekenmodule van het softwarepakket AERIUS¹⁶ dat bestuursorganen verplicht zijn te gebruiken. Met dit instrument kunnen bestuursorganen en burgers vaststellen hoeveel stikstofdepositie een activiteit veroorzaakt en welke ruimte per gebied bestaat om ‘vervuilende’ activiteiten toe te staan. Met de resultaten uit AERIUS Calculator wordt beoordeeld of een activiteit de drempel- of grenswaarde overschrijdt waardoor voor de activiteit een meldings- dan wel een vergunningsplicht bestaat en of een vergunning dan kan worden verleend.

Een groot voordeel van het PAS is de bijdrage die het levert aan de rationaliteit van de beslissing: het in samenhang beoordelen van maatregelen, gebaseerd op de beste wetenschappelijke kennis, kan de kwaliteit van de besluitvorming verhogen. Daarnaast geeft het efficiëntievoordelen, want snel kunnen maatregelen en cumulatieve gevolgen worden afgewogen en weet de burger waar hij aan toe is. Dit efficiëntievoordeel erkent ook de Afdeling.¹⁷ Het risico is alleen dat dergelijke deels geautomatiseerde besluitvorming niet inzichtelijk en controleerbaar is vanwege een gebrek aan inzicht in de gemaakte keuzes en gebruikte gegevens en aannames, aldus de Afdeling.¹⁸ Door dit ‘black box’-karakter kan een ongelijkwaardige procespositie ontstaan als belanghebbenden willen opkomen tegen een besluit: zij kunnen niet controleren op basis waarvan een besluit is genomen én of inderdaad de zekerheid bestaat dat de natuurgebieden niet zullen worden aangetast.

Om reële rechtsbescherming mogelijk te maken, legt de Afdeling een verplichting op aan de bestuursorganen die verantwoordelijk zijn voor de ontwikkeling van het programma en de software. Zij moeten de gemaakte keuzes en de gebruikte gegevens en aannames volledig, tijdig en uit eigen beweging openbaar maken op een passende wijze, zodat deze voor derden toegankelijk zijn.¹⁹ Daarmee kunnen belanghebbenden dergelijke besluiten (laten) controleren en eventueel gemotiveerd betwisten en kan de rech-

12 Artikel 6 lid 3 Richtlijn 92/43/EEG van de Raad van 21 mei 1992.

13 Zie ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, r.o. 10.11-10.33 met verwijzingen naar uitspraken van het HvJ EU en de interpretatie daarvan door de Afdeling.

14 Ondertussen is dit de minister van Infrastructuur en Waterstaat.

15 HvJ EU 7 november 2018, ECLI:EU:C:2018:882, c-293/17 en c-294/17, r.o. 101, AB 2018/436, m.nt. Ch. W. Backes. Zie ook: S.D.P. Kole, ‘PASst het PAS nog?’, *Gst.* 2019/3.

16 Te raadplegen op aerius.nl.

17 ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, r.o. 14.3.

18 ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, r.o. 14.3.

19 ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, r.o. 14.4.

ter deze toetsen op rechtmatigheid. Overigens is in een latere Afdelingsuitspraak de nuance aangebracht dat alleen de keuzes ten aanzien van de gegevens die de gebruiker zelf heeft ingevoerd, de *maatwerk invoergegevens*, uit eigen beweging moeten worden ontsloten.²⁰ De standaardinvoergegevens in de rekenmodule hoeven slechts op verzoek van belanghebbenden kenbaar te worden gemaakt om onevenredig zware bestuurslasten te voorkomen. Voor de ministers en staatssecretaris bleek er in dit geval werk aan de winkel. Hoewel AERIUS voor iedereen gratis toegankelijk was²¹ en er een uitleg van 190 factsheets beschikbaar was, oordeelde de Afdeling na grondige studie dat verschillende keuzes, gegevens en aannames onvoldoende onderbouwd waren. Het is nog afwachten of het de ministers en staatssecretaris lukt de kwaliteit en controleerbaarheid van het PAS voldoende te verbeteren.

De Afdeling legt met dit toetsingskader voor het algoritmische beslissysteem de nadruk op de ongelijkwaardige procespositie van belanghebbenden en de uit het beginsel van *equality of arms* vloeiende noodzaak voor bestuursorganen die de software ontwikkelden om de ingevoerde uitgangspunten toe te lichten en te onderbouwen. Dat ziet niet zozeer op de technische kant van het algoritme zelf, maar vooral op een toelichting van de gegevens, aannames en keuzes waarvan het systeem uitgaat, zoals een gemiddelde economische groei van 2,5% of van een doorzettende autonome daling van de gemiddelde stikstofdepositie.

2.3 De modelmatige waardebeoordeling in de WOZ-beschikking

Dit toetsingskader strekt zich ook uit naar andere (deels) geautomatiseerde besluitvorming, zoals de vaststelling van de WOZ-waarde van een onroerende zaak op basis van de Wet waardering onroerende zaken.²² Het gaat bij deze waarde om de prijs die de meestbiedende koper zou betalen als de woning aangeboden wordt na de beste voorbereiding en op de meest geschikte wijze.²³ Een heffingsambtenaar van de gemeente bepaalt de WOZ-waarde door modelmatige waardebeoordeling: een methode waarbij de waarde wordt getaxeerd door de kenmerken van en rondom de woning systematisch te vergelijken met de kenmerken en de daaraan gekoppelde waarden van recente woningverkoop in de gemeente.²⁴ Het is een computerprogramma dat deze kenmerken en waarden genereert en de woning daarmee vergelijkt. Gemeenten mogen zelf kiezen welk taxatiesysteem en programma het beste past bij de woningen binnen de gemeente.²⁵

De woningeigenaar die de opbouw van de WOZ-waarde wil controleren, kan een motivering van de WOZ-waarde vinden via MijnOverheid of een digitaal loket van de

20 ABRvS 18 juli 2018, ECLI:NL:RVS:2018:2454, r.o. 23, *Computerrecht* 2018/253, m.nt. N. Jak & T. Barkhuysen, *Me&R* 2018/124, m.nt. M.M. Kaajan.

21 In tegenstelling tot andere gebruikte software in het omgevingsrecht, zie R. Benhadi, 'Geautomatiseerde besluitvorming in het omgevingsrecht', *SLAB* 2017, afl. 3, p. 10.

22 HR 17 augustus 2018, ECLI:NL:HR:2018:1316, *NTFR* 2018/1878, m.nt. M.B. Weijers, *AB* 2018/388, m.nt. P.A. Flutsch.

23 *Kamerstukken II* 1993/94, 22 885, nr. 36, p. 44.

24 www.waarderingskamer.nl/klopt-mijn-woz-waarde/totstandkoming-woz-waarde/.

25 Notitie modelmatige waardebeoordeling, Waarderingskamer 21 augustus 2017, p. 3, te raadplegen op: www.waarderingskamer.nl/klopt-mijn-woz-waarde/totstandkoming-woz-waarde/.

gemeente.²⁶ Deze motivering bestaat veelal uit het taxatieverslag met de kenmerken van de woning en de waarde van drie verkochte woningen uit de buurt, maar niet met de gegevens uit het computerprogramma waarmee is vergeleken. Dit verslag bevat dus niet *alle* gegevens die nodig zijn om de vastgestelde waarde te kunnen controleren.²⁷ Mocht een woningeigenaar - na bezwaar - tegen een WOZ-beschikking in beroep gaan dan is het in de praktijk gebruikelijk dat de heffingsambtenaar een taxateur inschakelt om een nieuw taxatierapport uit te brengen en richt de rechtsstrijd zich vervolgens op dit rapport in plaats van de resultaten van het computerprogramma.²⁸

De Hoge Raad deed op 17 augustus 2018 uitspraak over de mogelijkheid om de WOZ-waarde te controleren in een zaak waarin een woningeigenaar had verzocht om een overzicht van de opbouw van de kavelwaarde van zowel de eigen woning als de referentieobjecten (de 'grondstaffels').²⁹ Deze woningeigenaar kreeg in de bezwaarfase nul op het rekest, omdat het volgens de heffingsambtenaar technisch niet mogelijk was om grondstaffels te herleiden uit het softwareprogramma. Daar gaat het Hof niet mee akkoord: een heffingsambtenaar dient zijn administratie zodanig in te richten dat de noodzakelijke bouwstenen voor de waardebepaling daaruit op doelmatige wijze zijn te destilleren.³⁰ Volgens de Hoge Raad vormen de grondstaffels op de zaak betrekking hebbende stukken die al in de bezwaar- of beroepsfase ter inzage moeten worden gegeven.³¹ Vervolgens herhaalt de Hoge Raad het toetsingskader van de Afdeling en benadrukt dat de heffingsambtenaar om een ongelijkwaardige procespositie te voorkomen had moeten zorgdragen voor de inzichtelijkheid en controleerbaarheid van de keuzes, aannames en gegevens die leiden tot de grondstaffels.³²

Het is opvallend dat de heffingsambtenaar software gebuikt voor het primaire besluit, maar bij aanwending van rechtsmiddelen een deskundige taxateur inschakelt om de waarde te bepalen en onderbouwen. Dat is vanuit rationaliteit lastig te begrijpen: in beginsel zal een bestuursorgaan voor een feitenvaststellingsmethode kiezen die de waarheid dicht benadert en met beperkte uitvoeringskosten gepaard gaat. Het is vanuit bewijsrechtelijk perspectief bepaald ongebruikelijk om het primaire besluit te baseren op een meer nauwkeurige goedkope analyse (via de software), om vervolgens bij de verlengde besluitvorming over te stappen naar de methode die weliswaar meer inzichtelijk te motiveren is, maar waarbij de uitkomst veel afhankelijker is van de drie (mogelijk strategisch) gekozen referentieobjecten.³³ Inzichtelijkheid gaat dan voor rationaliteit en neemt de prikkel weg om de software te verbeteren.

26 Op basis van artikel 3:48 lid 1 en 2 Awb en artikel 40 lid 2 Wet WOZ hoeft de motivering slechts op verzoek te worden gegeven.

27 Zie hierover ook: J.C.K.W. Bartel, 'De WOZ-beschikking: het roer moet om', *WFR* 2018/203, p. 1402-1403.

28 Zie voor een kritische beschouwing van deze praktijk: Bartel 2018.

29 HR 17 augustus 2018, ECLI:NL:HR:2018:1316, *NTFR* 2018/1878, m.nt. M.B. Weijers, *AB* 2018/388, m.nt. P.A. Flutsch.

30 Hof's-Hertogenbosch 10 februari 2017, ECLI:NL:GHSHE:2017:501.

31 Op grond van artikel 7:4 lid 2 en 8:42 lid 1 Awb.

32 HR 17 augustus 2018, ECLI:NL:HR:2018:1316, r.o. 2.3.3, *NTFR* 2018/1878, m.nt. M.B. Weijers, *AB* 2018/388, m.nt. P.A. Flutsch.

33 Vgl. Bartel 2018, p. 1406; de drie referentieobjecten lijken steeds te worden gekozen uit de groep van verkopen met bovengemiddelde waarden voor de objectkenmerken, terwijl het computerprogramma

3 Mens en systeem samen deskundig

3.1 Ondersteunende software

Waar in de voorgaande paragraaf het softwareprogramma een beslissende rol speelt in de besluitvorming, is in veel andere gevallen de software meer ondersteunend van aard. Aan een tekstverwerkingsprogramma zitten weinig haken en ogen, maar dat wordt anders wanneer met behulp van software beoordelingen in punten worden geduid en scores worden gestandaardiseerd, zoals bijvoorbeeld bij subsidietenders.³⁴ Is er een duidelijke handleiding voor welke standaard een bepaalde score symbool staat? Vindt bij een gelijke score een nadere inhoudelijke beoordeling plaats of wordt een applicatie gebruikt om *random* te selecteren en hoe is de *randomness* geborgd?³⁵ Wie kennis heeft van de gebruikte software vindt dit logische kritische kanttekeningen, wie niet weet welk programma is gebruikt, zal dit niet snel problematiseren. De casus waarin de ondersteunende software nadrukkelijk naar voren komt en door de bestuursrechter is getoetst, is de arbeidsongeschiktheidsbeoordeling door het Uitvoeringsinstituut Werknemersverzekeringen (Uwv) met behulp van het Claimbeoordelings- en borgingssysteem (CBBS).

3.2 Het CBBS

Bij een arbeidsongeschiktheidsbeoordeling brengt een verzekeringsarts de beperkingen en mate van arbeidsongeschiktheid in kaart, waarna een arbeidsdeskundige beziet welke functies betrokkene nog zou kunnen uitoefenen en wat zijn of haar resterende verdien capaciteit is.³⁶ Zij gebruiken daarbij een computerprogramma, het CBBS. Dit systeem heeft in wezen drie functies: het dient als communicatiemiddel tussen verzekeringsarts en arbeidsdeskundige, als instrument om (geautomatiseerd) een voorselectie van functies te maken én als middel om 'naar buiten toe' inzichtelijk te maken hoe tot het deskundige arbeidsongeschiktheidsoordeel is gekomen.³⁷ Nadat in 2002 dat programma het oude Functie Informatie Systeem (FIS) verving, zijn er verschillende rechtszaken gevoerd over de deugdelijkheid van arbeidsongeschiktheidsbeoordelingen met behulp van het CBBS.³⁸ Eisers meenden dat het programma onnavolgbaar was en van verkeerde uitgangspunten uitging.

In het ontwerp van het CBBS liggen namelijk verschillende aannames besloten.³⁹ Een belangrijk onderdeel van het CBBS is de Functionele Mogelijkheden Lijst (FML) waarmee de verzekeringsarts iemands belastbaarheid en beperkingen in kaart brengt. De

gegevens uit het gehele verkoopbestand genereert, waaronder ook die met beneden gemiddelde waarden.

34 Zie bijv. Rb. Den Haag 10 januari 2019, ECLI:NL:RBDHA:2019:141, over de standaardisatie van beoordelingsgemiddelden bij de verdeling van NWO-subsidies.

35 Zie bijv. CbB 23 augustus 2011, AB 2011/329, m.nt. Wolsinkel.

36 Art. 2, eerste lid, Schattingsbesluit arbeidsongeschiktheidswetten.

37 CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9971, USZ 2006/309, m.nt. M. Koolhoven.

38 CRvB 9 november 2004, ECLI:NL:CRVB:2004:AR4719, AB 2004/462, m.nt. F.J.L. Pennings, USZ 2004/353, m.nt. P.S. Fluit; CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9971, USZ 2006/309, m.nt. M. Koolhoven; CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9976, SR 2006/81, m.nt. A.C. Damsteegt.

39 De schets is ontleend aan CRvB 9 november 2004, ECLI:NL:CRVB:2004:AR4719, AB 2004/462, m.nt. F.J.L. Pennings, USZ 2004/353, m.nt. P.S. Fluit.

FML bevat een zestal rubrieken, waaronder ‘sociaal functioneren’ en ‘statische houdingen’, met daarin belastbaarheidsaspecten die de verzekeringsarts moet scoren. De FML neemt daarbij bepaalde normaalwaarden tot ijkpunt, die aangeven tot welk functioneren een gezond persoon tussen de 16 en 65 jaar minimaal in staat is. Die normaalwaarden zijn gebaseerd op professionele consensus van een panel van arbeidsdeskundigen en verzekeringsartsen, de beschikbare informatie over het gebruik van het FIS en adviezen van wetenschappelijke specialisten.⁴⁰ De verzekeringsarts geeft per afzonderlijk belastbaarheidsaspect aan of iemand ten opzichte van de normaalwaarde beperkt is te achten, waarbij hij vier gradaties onderscheidt. Vervolgens moet een arbeidsdeskundige beoordelen welke functies betrokkene nog kan vervullen. Hiervoor zijn in het CBBS allerhande functies ingevoerd en geclusterd op gelijksoortigheid.⁴¹ Het programma vergelijkt de belastbaarheidsaspecten uit de FML met in het systeem opgenomen belastinggegevens voor de functies en selecteert automatisch potentieel geschikte functies op basis van ‘matchende’ beoordelingspunten. De geautomatiseerde matching is niet in alle opzichten passend. Sommige belastbaarheidsaspecten in de FML corresponderen niet met gegevens over de functies, omdat deze aspecten niet ‘in maat en getal’⁴² zijn uit te drukken en per werksituatie verschillen. Als gevolg zal de arbeidsdeskundige, eventueel in overleg met de verzekeringsarts, de functieselectie handmatig moeten bijwerken.

Een ondersteunend systeem als het CBBS, dat standaardiseert en selecteert op basis van aannames, is volgens de Centrale Raad van Beroep toelaatbaar: “De Raad is niet gebleken van redenen om een systeem als het CBBS niet in beginsel rechtens aanvaardbaar te achten als ondersteunend systeem, omdat wet en rechtspraak onvoldoende aanknopingspunten bieden om daarover in algemene zin anders te oordelen, waarbij hij aantekent dat het UWV enige beoordelingsruimte niet kan worden ontegd.”⁴³ Met deze redenering lijken weinig uitvoeringstrajecten zich te verzetten tegen de inzet van ondersteunende software. De Raad accepteert ook de gebruikte normaalwaarden in het systeem, maar hecht daarbij expliciet waarde aan de gebruiksinstructie waarin mogelijkheden worden genoemd om handmatig aanpassingen te doen in de functieselectie of bij belastbaarheidsaspecten aanvullende voorwaarden te stellen.

Wel achtte de CRvB de eerste versie van het CBBS gebrekkig omdat de wijze van totstandkoming van de schatting minder inzichtelijk, verifieerbaar en toetsbaar was dan bij een schatting met behulp van het oude systeem, het FIS. Interessant is dat het CBBS juist is ontwikkeld als reactie op het FIS, waarin een verzekeringsarts op 28 aspecten de mate van beperking van betrokkene moest aangeven. Daarop kwam vanuit verzekeringsartsen de kritiek dat dit systeem veel meer objectiviteit suggereerde dan (professioneel) verantwoord was en waardoor onverklaarbaar grote afwijkingen ontstonden in de oordelen van artsen.⁴⁴ Het CBBS moest meer recht doen aan de complexe beoordeling. Waar het tot meer nuance moest leiden, is het resultaat dat het programma voor de

40 Bijlage bij *Kamerstukken II* 2005/06, 28 333, nr. 72.

41 Deze zijn herkenbaar aan een SBC-code (Standaard Beroepen Classificatie).

42 Bijlage bij *Kamerstukken II* 2005/06, 28 333, nr. 72.

43 CRvB 9 november 2004, ECLI:NL:CRVB:2004:AR4719, AB 2004/462, m.nt. F.J.L. Pennings, *USZ* 2004/353, m.nt. P.S. Fluit.

44 Zie de noot van Pennings, aant. 1 onder CRvB 9 november 2004, ECLI:NL:CRVB:2004:AR4719, AB 2004/462.

burger en de rechter minder transparant is. Het UWV werd opgedragen om binnen een bepaalde termijn het CBBS aan te passen om het systeem voor zowel de betrokkene als de rechter beter controleerbaar te maken.⁴⁵ Hoewel de Centrale Raad weinig expliciet is over de bestuursrechtelijke normen waaruit deze bezwaren voortvloeien, hanteert hij het zorgvuldigheids- en motiveringsbeginsel als vernietigingsgrond.

De Centrale Raad lijkt hiermee vooral kritisch te zijn op de derde functie van het CBBS om 'naar buiten toe' inzichtelijk te maken waar het deskundige arbeidsongeschiktheidsoordeel op is gebaseerd.⁴⁶ Ook in latere rechtszaken staat de motivering van op basis van het CBBS gegeven oordelen centraal. Opvallend is dat rechtbanken in eerste aanleg soms nog verder wilden gaan in het verantwoorden van het beslissysteem, zoals de motiveringseis om uit te leggen hoe frequent een bepaalde handeling binnen de normaalwaarden kan worden verricht.⁴⁷ Zo ver wil de Raad niet gaan. Het CBBS is uiteindelijk niet meer dan een hulpmiddel. Het is de arbeidsdeskundige die, eventueel na overleg met de verzekeringsarts, steeds tot een toereikend gemotiveerd eindoordeel moet komen.⁴⁸

Wat opvalt aan deze casus is dat vanuit de medische expertise een nieuw ondersteunend systeem wenselijk werd geacht, omdat het oude systeem een overnauwkeurigheid vroeg die professioneel niet verantwoord was. Het FIS scoorde daarom laag op rationaliteit, maar hoog op controleerbaarheid. Het CBBS daarentegen versterkt de rationaliteit, omdat het vanuit de verzekeringsgeneeskundige en arbeidskundige kant methodisch bruikbaar is dan het FIS (en bijvoorbeeld meer recht doet aan psychische problematiek). Ook genereert het CBBS meer informatie over de verschillen tussen beoordelingen per regio, kantoor of medewerker.⁴⁹ Tegelijkertijd boet het CBBS in aan controleerbaarheid van de individuele beoordeling, omdat de geautomatiseerde analyse nog handmatige bewerking vergt.

De controleerbaarheid van de arbeidsongeschiktheidsbeoordelingen blijft tot debat bij de rechter leiden, vooral op basis van het beginsel van *equality of arms*. Regelmatig betogen betrokkenen dat daar strijd mee is, omdat zij niet in staat zijn het gebruikte systeem danwel andere aspecten van deskundigenadvisering te betwisten. De Centrale Raad wijst echter de grond van de hand dat betrokkene over alle data uit het CBBS systeem zou moeten kunnen beschikken of toegang zou moeten krijgen tot het systeem zelf.⁵⁰ Wel moet het Uwv, vergelijkbaar met de PAS-uitspraak, bij een gemotiveerde

45 CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9971, *USZ* 2006/309, m.nt. M. Koolhoven; CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9976, *SR* 2006/81, m.nt. A.C. Damsteegt. Zo moest er een voor belanghebbende en rechter zichtbare signalering komen bij geselecteerde functies die mogelijk de belastbaarheid van betrokkene overstegen.

46 Zie hierover ook de annotatie van Van Eck bij ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, *Computerrecht* 2017/256, m.nt. B.M.A. van Eck.

47 Zie de annotaties bij CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9971, *USZ* 2006/309, m.nt. M. Koolhoven; CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9976, *SR* 2006/81, m.nt. A.C. Damsteegt.

48 CRvB 12 oktober 2006, ECLI:NL:CRVB:2006:AY9971, *USZ* 2006/309, m.nt. M. Koolhoven. Zie ook CRvB 1 februari 2008, ECLI:NL:CRVB:2008:BC3237, *RSV* 2008/52.

49 Bijlage bij *Kamerstukken II* 2005/06, 28 333 nr. 72.

50 CRvB 10 januari 2019, ECLI:NL:CRVB:2019:183; CRvB 13 februari 2019, ECLI:NL:CRVB:2019:606.

beroepsgrond de ‘standaardinvoergegevens’ van het systeem verstrekken, zoals gegevens over de actualisering van het functiebestand.⁵¹

4 De deskundige mens

4.1 Deskundigenadvisering

In tegenstelling tot het jonge toetsingskader voor geautomatiseerde besluitvorming bestaat voor deskundigenadvisering al jaren min of meer hetzelfde stramien. Het algemene toetsingskader richt zich op de normen van zorgvuldigheid, motivering, onpartijdigheid en steeds meer op *equality of arms*. De kern van de toetsing door de bestuursrechter ligt in de vraag of het bestuursorgaan zijn vergewisplicht heeft vervuld.⁵² Het bestuursorgaan dient te controleren of de deskundige onpartijdig en voldoende deskundig is, of het deskundigenadvies niet gebrekkig tot stand is gekomen en - tot op zekere hoogte - of de inhoud van het advies deugdelijk is.⁵³ Daarnaast is het beginsel van *equality of arms* onder artikel 6 EVRM een steeds nadrukkelijker stempel op de beoordeling gaan drukken. Bij deskundigenbewijs betekent dit onder andere dat partijen de mogelijkheid moeten hebben om effectief op deskundigenrapporten te reageren, eventueel met behulp van een tegenrapport, als een deskundigenadvies een zwaarwegende invloed heeft op de beslissing van de rechter.⁵⁴

We bezien in twee casus hoe dit zeer globale toetsingskader uitwerkt en zich verhoudt tot waarden als rationaliteit en accountability. We hebben gekozen voor planschadetaxaties (in 4.2), omdat deze adviezen grote invloed hebben op het besluitvormingsproces over tegemoetkomingen in planschade,⁵⁵ de vaststelling enige parallellen vertoont met de WOZ-beschikking en de bestuursrechter hier expliciet erkent dat een taxatie gedeeltelijk subjectief is. Vervolgens gaan we in op de medische screening bij asielzaken (in 4.3). Daar speelt de advisering een meer ondersteunende rol voor het kunnen waarderen van de mondelinge verklaringen en is de aangewende expertise meer organisatorisch ingebed. Voor beide casus geldt dat de Afdeling de eisen recent heeft aangescherpt.

4.2 Planschadetaxaties bij onroerende zaken

In tegenstelling tot de taxatie van de WOZ-waarde komen planschadetaxaties niet geautomatiseerd tot stand, maar (enkel) met behulp van een deskundige. Bij planschade draait het om de vraag of, en zo ja in hoeverre, een aanvrager bij een wijziging van het

51 CRvB 10 januari 2019, ECLI:NL:CRVB:2019:183.

52 Artikel 3:9 Awb of afgeleid uit artikel 3:2 Awb. Zie ook M. Schreuder-Vlasblom, *Rechtsbescherming en bestuurlijke voorprocedure*, Deventer: Kluwer 2017, p. 718.

53 Zie hiervoor recent: I.S. Ouweland, 'Deskundigenadvisering in de bestuurlijke fase - het toetsingskader van de bestuursrechter onder de loep', *www.rechtenoverheid.nl* 2019, p. 1-53; L.M. Koenraad, 'Deskundigen in het Nederlands bestuursrecht. Rechterlijke controle van bestuursoptreden dat op deskundigenadvisering is gebaseerd', in: I. Opdebeek e.a., *Vereniging voor de vergelijkende studie van het recht van België en Nederland - preadviezen 2017*, Den Haag: Boom Juridisch 2017, p. 93-98.

54 Zie bijvoorbeeld EHRM 18 maart 1997, ECLI:NL:XX:1997:AD4449 (Mantovanelli), *NJ* 1998/278, m.nt. H.J. Snijders.

55 G.M van den Broek & M.K.G. Tjepkema, *De reikwijdte en rechtsgrondslag van nadeelcompensatie in het omgevingsrecht (preadviezen voor de Vereniging voor Bouwrecht)*, nr. 43, 2015, p. 86.

planologische regime recht heeft op een tegemoetkoming voor bijvoorbeeld de waardevermindering van een woning. Bij de beoordeling van aanvragen is het gebruikelijk dat het bestuursorgaan een externe deskundige inschakelt.⁵⁶ De ratio hierachter is dat advisering enerzijds nodig is om het besluit zorgvuldig voor te bereiden als het bestuursorgaan zelf de specifieke deskundigheid mist en anderzijds om te faciliteren dat de aanvrager eenvoudig toegang heeft tot een onafhankelijke deskundige.⁵⁷

De taak van de deskundige bestaat onder andere uit de taxatie van de omvang van de schade.⁵⁸ De maatstaf daarvoor is welke prijs een redelijk denkend en handelend koper voor de onroerende zaak zou hebben geboden onmiddellijk voor de inwerkingtreding van het nieuwe planologische regime en op het tijdstip direct daarna, uitgaande van wat maximaal op grond van het oude en nieuwe planologische regime kan worden gerealiseerd.⁵⁹ Verschillende kenmerken van het planschaderecht maken deze taxatie complex.⁶⁰ Zo is de feitelijke situatie in beginsel niet van belang, maar wordt uitgegaan van een vergelijking van wat *maximaal* op grond van het oude en het nieuwe planologische regime kan worden gerealiseerd, ongeacht of verwezenlijking heeft plaatsgevonden.⁶¹ Dit kenmerk is één van de belangrijkste redenen waarom een waardedaling op grond van de WOZ niet volstaat om de omvang van de planschade vast te stellen, want daar is vooral de feitelijke situatie leidend.⁶² Een ander complex element is dat de waardebepalingen moeten worden geobjectiveerd en gerelativeerd. Alleen de door de planologische maatregel veroorzaakte verslechtingen worden meegewogen (zoals verlies van uitzicht of lichtinval) en dan alleen voor zover deze verslechtingen zien op objectief te verwachten, ruimtelijke gevolgen van het planologische regime.⁶³ Subjectieve elementen als negatieve gevoelswaarden of onbestemde angsten spelen geen rol.⁶⁴ Deze complexiteit heeft geleid tot de kritiek dat de tegemoetkoming in planschade 'is losgezongen van de ervaringswereld van burger en overheid'⁶⁵ en is een belangrijke stimulans om in de

56 Dat is vooralsnog wettelijk verplicht, art. 6.1.3.2 Besluit ruimtelijke ordening. In de toekomstige Omgevingswet lijkt dit niet meer het geval. Zie ABRvS 25 oktober 2017, ECLI:NL:RVS:2017:2893 (*Helmond II*), *Gst.* 2018/32, m.nt. P.C.M. Heinen, aant. 5 en T. ten Have & J.J. Thoonen, 'De rol van de externe deskundige bij nadeelcompensatie in de Omgevingswet', *O&A* 2017/66.

57 B.P.M. van Ravens, 'Over de onafhankelijke deskundige in procedures over overheidsaansprakelijkheid', in: T. Barkhuysen e.a. (red.), *25 jaar. In eenheid en verscheidenheid*, Deventer: Kluwer 2019, p. 186-188.

58 Artikel 6.1.3.4 Besluit ruimtelijke ordening.

59 ABRvS 28 september 2016, ECLI:NL:RVS:2016:2582 (hierna: overzichtsuitspraak), r.o. 4.6. Zie voor een mogelijk andere lijn in de jurisprudentie van de Afdeling: ABRvS 14 juni 2017, ECLI:NL:RVS:2017:1580 en ECLI:NL:RVS:2017:1579, *BR* 2017/75 en *BR* 2017/76, m.nt. J.W. van Zundert.

60 B.P.M. van Ravens, 'Het is eigenlijk heel eenvoudig... Abstraheren, maximaliseren, objectiveren, relateren, toerekenen en taxeren bij het beoordelen van planschade', in: A.A.J. de Gier & G.T.M. Jurgens (red.), *Goed verdedigbaar. Vernieuwing van het bestuursrecht*, Deventer: Kluwer 2011, p. 131-148 en S. Berns, 'De planschadetaxatie', *PRO* 2016/84.

61 Overzichtsuitspraak planschade, r.o. 2.1.

62 N.A. Boerties, 'Waardevermindering in het kader van planschade en WOZ: ééneige tweelingen of verre familie?', *Praktijk omgevingsrecht* 2017, afl. 2, p. 60-64. Zie nog recent een bevestiging in: ABRvS 20 maart 2019, ECLI:NL:RVS:2019:870, r.o. 6.

63 Overzichtsuitspraak planschade, r.o. 2.11 en 2.12.

64 Zelfs al is schade aantoonbaar, zie bijvoorbeeld ABRvS 5 juli 2017, ECLI:NL:RVS:2017:1752 en Overzichtsuitspraak planschade, r.o. 2.12.

65 Van Ravens 2011, p. 132.

Omgevingswet te zoeken naar een systeem dat meer aansluit bij wat in werkelijkheid verandert in de fysieke leefomgeving.⁶⁶

Belanghebbenden gaan regelmatig tegen planschadebesluiten in beroep, waarbij zowel gronden tegen de taxatiemethode als tegen de persoon van de deskundige worden gericht. Deskundigenadviezen kunnen fors van elkaar verschillen door de grote mate van 'vrijheid' die de deskundige heeft om op basis van kennis, ervaring en intuïtie taxatiemethodes te kiezen en in te schatten welke waardedaling reëel is. De meest gebruikelijke methode voor het taxeren van woningen is de vergelijkingsmethode, waarbij de deskundige de geschatte waarde vóór de planologische wijziging vergelijkt met die van ná de planologische wijziging aan de hand van referentieobjecten in de markt.⁶⁷ Er zijn echter ook alternatieve methoden, zoals de (vanwege zijn schijnbare objectiviteit omstreden) schadefactorenmethode. Deze methode komt erop neer dat vanuit één geschatte waarde de andere waarde wordt beredeneerd door schadefactoren te wegen en deze weging te koppelen aan een waarderingsverminderingspercentage.⁶⁸ Zowel de keuze voor een bepaalde methode als de toepassing daarvan kan leiden tot een verschillende omvang van de schade, wat willekeurig kan overkomen op de burger.

Hoe toetst de Afdeling deze taxaties? De standaardoverweging van de Afdeling is dat het bestuursorgaan van een advies mag uitgaan als daarin door een onafhankelijke en onpartijdige deskundige op objectieve wijze verslag is gedaan van het door deze deskundige verrichte onderzoek en op inzichtelijke wijze is aangegeven welke feiten en omstandigheden aan de conclusie ten grondslag zijn gelegd en deze conclusies niet onbegrijpelijk zijn, tenzij concrete aanknopingspunten voor twijfel aan de juistheid of volledigheid ervan naar voren zijn gebracht.⁶⁹ De Afdeling benadrukt dat de waardering van onroerende zaken niet slechts door het toepassen van een taxatiemethode plaatsvindt, maar dat daarbij ook kennis, ervaring en intuïtie van de desbetreffende deskundige een rol speelt.⁷⁰ Deze schattingen vragen zorgvuldig maatwerk, afgestemd op lokale omstandigheden en marktverhoudingen, wat moeilijk in een (algoritmische) formule te vangen is.⁷¹ Het gaat erom dat de deskundige binnen een bepaalde bandbreedte is gebleven en het bestuursorgaan zich in redelijkheid op het advies heeft kunnen baseren.⁷² Wel is de Afdeling strengere eisen gaan stellen aan de persoon van de deskundige, waarbij de schijn van partijdigheid moet worden voorkomen.⁷³

66 *Kamerstukken II* 2017/18, 34 986, nr. 3, p. 24 (MvT Invoeringswet Omgevingswet).

67 Bijvoorbeeld P.A.H.M. Willems, 'Planschade: een vermindering van de waarde van een onroerende zaak (2)', *StAB* 2016, afl. 4, p. 9 en S. Berns, 'De planschadetaxatie', *PRO* 2016/84.

68 Zie discussie over deze methode in: A. Nentjes, 'Over de juiste raming van bestemmingsplanschade', *NJB* 2008/1857; B.P.M. van Ravens, 'Schadevergoeding (w.o. onrechtmatige overheidsdaad)', *NTB* 2009/4, p. 23-24; A. Nentjes, 'Planschade: het adviesbureau, de taxatiedeskundige en de rechter', *NTB* 2010/6; L.A. van Montfoort, 'Planschade; adviesbureaus en taxatiemethoden', *NTB* 2010/17, p. 93-95.

69 Overzichtsuitspraak planschade, r.o. 8.3.

70 Overzichtsuitspraak planschade, r.o. 8.11.

71 L.A. van Montfoort, 'Planschade; adviesbureaus en taxatiemethoden', *NTB* 2010/17, p. 95.

72 ABRvS 14 november 2012, ECLI:NL:RVS:2012:BY3052; overzichtsuitspraak planschade, r.o. 8.11.

73 Overzichtsuitspraak planschade, r.o. 8.2; ABRvS 13 juli 2016, ECLI:NL:RVS:2016:1970, *AB* 2016/454, m.nt. J.H.M. Huijts.

De planschadetaxatie scoort relatief laag op rationaliteit en op inzichtelijkheid en controleerbaarheid van het advies. De inschatting van de schade kan sterk verschillen naar gelang de gekozen methode en deskundige, wier kennis, ervaring en intuïtie een rol mogen spelen. Noch de methode, noch de deskundigheid zijn uitgebreid ingebed in wetenschappelijke analyses of normen vanuit een beroepsgroep. Een planschadetaxateur is geen beschermde titel en kwalificatie-eisen ontbreken.⁷⁴ Waar we een dergelijke lage score op rationaliteit en inzichtelijkheid van een algoritmisch beslissysteem waarschijnlijk niet zouden accepteren, doen we dat wel in deze casus bij de adviserende mens. Wel eist de Afdeling dat de gedachtegang van de taxateur duidelijk en voldoende controleerbaar is.⁷⁵ Dit uit zich in motiveringseisen. Zo bepaalde de Afdeling dat het verschil tussen de WOZ-waarde en de planschadetaxatie moet worden uitgelegd, als dit verschil aanzienlijk is.⁷⁶ Ook moet bijvoorbeeld de gebruikte methode worden onderbouwd en moeten, in het geval van de vergelijkingsmethode, de referentieobjecten genoemd (kunnen) worden.⁷⁷ Het beginsel van *equality of arms* krijgt minder nadruk en het leveren van tegenbewijs is niet eenvoudig, zeker omdat een contra-expertise met een afwijkende waardebepaling door de subjectieve elementen niet als vanzelf tot de conclusie leidt dat het door het bestuursorgaan ingewonnen advies gebrekkig was.⁷⁸

4.3 Medische screening voor het gehoor in asielzaken

Waar de planschadetaxatie een vrij beroep is, is het artsenvak dat allerminst. In verschillende besluitvormingsregimes is medische expertise nodig, waarvoor een arts verantwoordelijkheid moet nemen. De medische deskundige is bij uitstek de expert voor wie strenge registratie- en opleidingseisen gelden, die zich moet houden aan de normen van de beroepsgroep op straffe van tuchtrechtelijke maatregelen en wier vakgebied streeft naar rationaliteit en *evidence-based medicine*. Achter de adviserende mens zit een stevig 'systeem'. Zorgt dat voor aan andere toetsingswijze door de bestuursrechter? Wij geven een eerste aanzet met een casus naar de medische screening voor het gehoor in asielzaken.

In een asielprocedure moet de IND namens de staatssecretaris beslissen op een aanvraag voor een verblijfsvergunning.⁷⁹ Deze besluiten leunen sterk op wat de vreemdeling zelf over zijn of haar situatie verklaart, omdat *objectief* bewijsmateriaal over de herkomst, achtergrond, identiteit en vluchtmotieven van de asielzoeker veelal ontbreekt.⁸⁰ IND-beslisambtenaren waarderen het beschikbare bewijsmateriaal en beoordelen of de verklaringen van de asielzoeker geloofwaardig zijn. Een belangrijke voorwaarde voor ge-

74 B.P.M. Van Ravels, 'Deskundigenadvisering bij nadeelcompensatie en tegemoetkoming in planschade', *O&A* 2015/88, p. 165-167.

75 Zie bijvoorbeeld recent: ABRvS 6 maart 2019, ECLI:NL:RVS:2019:720, r.o. 9.2.

76 Zie bijv. ABRvS 16 juni 2010, ECLI:NL:RVS:2010:BM7718 en ABRvS 24 mei 2017, ECLI:NL:RVS:2017:1402.

77 Zie ABRvS 12 maart 2008, ECLI:NL:RVS:2008:BC6374, r.o. 2.4.1 en J.W. van Zundert, T&C Ruimtelijk bestuursrecht, commentaar op artikel 6.1 Wet ruimtelijke ordening.

78 Zie bijvoorbeeld ABRvS 23 oktober 2013, ECLI:NL:RVS:2013:1663, r.o. 8.1.

79 Zie de rechtsgronden voor verlening in artikel 29 Vw 2000.

80 Adviescommissie voor Vreemdelingenzaken, *De geloofwaardigheid gewogen. Een advies over het onderzoeken, integraal beoordelen en toetsen van verklaringen in de asielprocedure*, Den Haag: ACVZ 2016, p. 21.

loofwaardigheid is dat de verklaringen van de asielzoeker samenhangend en aannemelijk worden bevonden en niet in strijd zijn met beschikbare relevante informatie.⁸¹

Vanwege de invloed van deze verklaringen op het besluit, is van groot belang om vroegtijdig te onderzoeken of de asielzoeker in staat is om coherent te verklaren of dat medische of psychische beperkingen dit belemmeren.⁸² Daarom is sinds 2010 een vast onderdeel van de asielprocedure dat de asielzoeker voorafgaand aan het gehoor de mogelijkheid krijgt zich medisch te laten onderzoeken.⁸³ Nadat tussen 2010 en 2015 MediFirst de medische screening uitvoerde, heeft in 2015 de Forensisch medische maatschappij Utrecht (FMMU) de opdracht voor deze screening gegund gekregen. Deze wisseling van de wacht is omstreden, omdat FMMU won door een betere score op prijs maar een lagere score op kwaliteit dan MediFirst.⁸⁴ De medische screening van FMMU resulteert in een ‘medisch advies horen en beslissen’ dat objectieve informatie⁸⁵ moet bevatten voor IND-medewerkers om rekening mee te houden tijdens het gehoor en bij de interpretatie ervan in de beslissing.⁸⁶

Volgens de Afdeling mag de staatssecretaris ervan uitgaan dat een FMMU-advies, dat is opgesteld conform het Protocol Medisch advies Horen en Beslissen (het protocol), voldoet aan de vanuit het oogpunt van vakkundigheid te stellen eisen.⁸⁷ Dit protocol beschrijft de werkwijze van FMMU en bevat scholings- en ervaringseisen voor het personeel. Een verpleegkundige doet eerst onderzoek naar de medische situatie van de asielzoeker aan de hand van een (gefixeerd) onderzoeksformulier.⁸⁸ Na dit onderzoek kan de verpleegkundige zelf een advies opstellen of de asielzoeker doorverwijzen naar een bij FMMU werkzame arts. De arts draagt de eindverantwoordelijkheid voor het onderzoek en het advies en zal het formulier en advies in alle gevallen moeten controleren en ondertekenen. Alleen het advies wordt vervolgens in het procesdossier gevoegd. Zowel de verpleegkundigen als de artsen moeten ingeschreven zijn in het BIG-register en jaarlijks deelnemen aan onderwijsmodules en actualiteitenbijeenkomsten.⁸⁹ De arts moet daarnaast ingeschreven zijn bij de Registratie Commissie Specialisten.

Als het advies conform dit protocol is opgesteld én het advies ook nog inzichtelijk en concludent is, heeft de staatssecretaris aan zijn vergewisplicht voldaan.⁹⁰ Het daarop

81 Artikel 31, zesde lid, onder c, Vw 2000 gebaseerd op artikel 4, vijfde lid, Definitierichtlijn (Richtlijn 2011/95/EU van het Europees Parlement en de Raad van 13 december 2011).

82 Zie over deze belemmeringen bijvoorbeeld E.J.J.M. Bloemen & K.M. Zwaan, “Medisch beperkt” in asiel. Medische en psychische problemen in de asielprocedure en de mogelijkheden van werkinstructie 2010/13’, *A&MR* 2011, afl. 9, p. 398-406.

83 Artikel 3.109, lid 6, Vreemdelingenbesluit 2000.

84 A.M. Reneman, *Identification of asylum seekers with special reception and procedural needs in the Dutch asylum procedure*, VU Migration Law Series no. 16, september 2018, p. 54. Zie ook de rechtszaken van MediFirst: Rb. Den Haag (voorzieningenrechter), 6 oktober 2014, ECLI:NL:RBDHA:2014:13392 en Hof Den Haag, 25 november 2014, ECLI:NL:GHDHA:2014:3758.

85 Volgens het protocol mag het advies geen subjectieve beoordelingen bevatten.

86 IND-werkinstructie nr. 2010/13.

87 ABRvS 28 mei 2015: ECLI:NL:RVS:2015:1783, r.o. 3.3, *AB* 2015/290, m.nt. R.W.J. Severijns en ABRvS 27 juni 2018, ECLI:NL:RVS:2018:2086, r.o. 8.1, *AB* 2018/373, m.nt. A.C. Hendriks.

88 ABRvS 27 juni 2018, ECLI:NL:RVS:2018:2086, r.o. 4, *AB* 2018/373, m.nt. A.C. Hendriks. Zie ook Reneman 2018, p. 67.

89 ABRvS 27 juni 2018, ECLI:NL:RVS:2018:2086, r.o. 4.1, *AB* 2018/373, m.nt. A.C. Hendriks.

90 ABRvS 28 mei 2015: ECLI:NL:RVS:2015:1783, r.o. 3.3, *AB* 2015/290, m.nt. R.W.J. Severijns.

gebaseerde besluit is dan in beginsel in overeenstemming met het zorgvuldigheids- en motiveringsbeginsel genomen. De inhoud van het FMMU-advies kan alleen worden betwist door middel van een andersluidend deskundigenadvies. In de praktijk komen dergelijke tegenonderzoeken regelmatig van het instituut Mensenrechten en Medisch Onderzoek (iMMO).

Tot voorkort accepteerde de Afdeling dat de staatssecretaris op dit punt geen waarde toekeende aan de rapporten van iMMO, omdat iMMO-onderzoeken enige tijd ná de gehoren worden verricht.⁹¹ Dit leidde tot onvrede, omdat tegenbewijs leveren op deze manier vrijwel onmogelijk was.⁹² Hoewel de Afdeling de spanning met het beginsel van *equality of arms* niet noemt, wijzigt zij in 2018 haar jurisprudentie: de staatssecretaris kan concludente en inzichtelijke iMMO-rapporten niet links laten liggen met het argument dat zij door het enkele tijdsverloop niet betrouwbaar zijn.⁹³ Ook stelt de Afdeling dat de staatssecretaris een medisch deskundige moet raadplegen - of op zijn minst moet motiveren waarom hij deze niet raadpleegt - als het iMMO-rapport concludeert dat psychische problematiek zeer waarschijnlijk het verklarende vermogen van de asielzoeker heeft beïnvloed, mits het rapport aan enkele voorwaarden voldoet.⁹⁴ Zo moet het rapport bijvoorbeeld expliciteren op welke onderdelen van het asielrelaas de beperking van het vermogen om consistent en coherent te verklaren, invloed heeft gehad.

De Afdeling laat zich in deze uitspraak opvallend uitgebreid voorlichten over de werkwijze van zowel FMMU als van het iMMO. Voor de benodigde vakkundigheid en zorgvuldigheid van FMMU hecht de Afdeling vooral aan de personeelseisen in het Protocol en de consistente werkwijze waarbij de arts steeds controleert.⁹⁵ De sterke inbedding van het medisch oordeel maakt dat gemakkelijk aan het zorgvuldigheid- en motiveringsbeginsel kan worden voldaan en dat het deskundigenoordeel in beginsel hoog scoort op rationaliteit. De Afdeling blijft echter weg van de toetsing van de *kwaliteit* van de adviezen, terwijl op de hierboven genoemde punten sinds de overgang van MediFirst naar FMMU wel kritiek is.⁹⁶ Het niveau van het personeel is lager dan bij MediFirst en naleving van de bijscholingsverplichtingen wordt niet gecheckt.⁹⁷ De onderzoeken door de verpleegkundige lijken korter dan bij MediFirst en aanzienlijk minder vaak wordt een arts erbij gehaald.⁹⁸ Bovendien zouden FMMU-adviezen vaak te oriënterend van aard zijn om de complexe problematiek op te sporen en te weinig instructief voor IND-medewerkers door 'vage' of subjectieve omschrijvingen.⁹⁹ De standaarden waaraan iMMO-rapporten daarentegen moet voldoen lijken hoger, omdat deze rapporten altijd moeten worden opgesteld door artsen en zij de asielzoeker gemiddeld vier uur onder-

91 ABRvS 28 mei 2015; ECLI:NL:RVS:2015:1783, r.o. 3.5, AB 2015/290, m.nt. R.W.J. Severijns.

92 Zie hierover uitgebreid: Reneman 2018, p. 15-80 en 194-196.

93 ABRvS 27 juni 2018, ECLI:NL:RVS:2018:2086, r.o. 8.5, AB 2018/373, m.nt. A.C. Hendriks, het tijdsverloop is niet doorslaggevend omdat de conclusies in het iMMO-rapport alleen worden gebaseerd op medische informatie uit de periode van de gehoren.

94 ABRvS 27 juni 2018, ECLI:NL:RVS:2018:2086, r.o. 8.5, AB 2018/373, m.nt. A.C. Hendriks.

95 ABRvS 27 juni 2018, ECLI:NL:RVS:2018:2086, r.o. 8.1, AB 2018/373, m.nt. A.C. Hendriks.

96 Zie Hendriks in zijn annotatie onder AB 2018/373, aant. 7.

97 Reneman 2018, p. 62.

98 Reneman 2018, p. 70.

99 E. Bloemen & M. Kollen, 'Bewijs geleverd. Medisch steunbewijs en medische beperkingen in asiel', JNVR 2015/27, p. 64; Reneman 2018, p. 77-80.

zoeken.¹⁰⁰ Dit illustreert dat niet alleen in een algoritmisch beslissysteem beleidsmatige keuzes worden ingebouwd die de inhoud van een waardeoordeel mede bepalen, maar ook bij menselijke advisering. Bij de overgang van MediFirst naar FMMU is rationaliteit afgewogen tegen uitvoeringskosten en snelheid. Overigens blijft de Afdeling niet helemaal weg bij de inhoudelijke kant van de aangewende expertise; zij verlangt dat de contra-expert specifiek duidt op welke verklaring de psychische problematiek invloed heeft gehad. Dat is voor een inzichtelijk juridisch debat met bewijs en tegenbewijs prettig, maar het is de vraag of de deskundige zo een specifiek oordeel vanuit zijn wetenschapsgebied voor zijn rekening kan nemen.¹⁰¹ Juridische eisen kunnen ook afbreuk doen aan rationaliteit.

5 Rationaliteit en accountability van de adviserende mens en machine

5.1 Toetsingskaders vergeleken

In de voorgaande paragrafen hebben we in enkele casus laten zien aan de hand van welke normen de bestuursrechter het deskundigenoordeel en het beslissysteem benadert. Het zorgvuldigheids- en motiveringbeginsel en het beginsel van gelijke processuele kansen krijgen de nadruk, ingevuld en toegepast op de specifieke kenmerken van elke casus. De gemene deler is dat door hulpmiddelen van deskundigenadvisering en beslissystemen een bestuursorgaan de rationaliteit van een beslissing kan verhogen,¹⁰² maar óf dat daadwerkelijk geschiedt voor de rechter bijzonder lastig valt vast te stellen. De bestuursrechter stapt dan al snel over naar meer procedurele waarborgen. In onze casus zien we bij deskundigenadvisering vooral aandacht voor de geprotocolleerde werkwijze en personeelseisen, als ook een voorzichtige poging om bewijslevering beter mogelijk te maken (bij FMMU), en aandacht voor onpartijdigheid (bij planschade). Bij de deskundige beslissystemen trekt de rechterlijke aandacht naar de inzichtelijkheid en controleerbaarheid van de gemaakte beoordeling. Vanuit dit perspectief is kruisbestuiving mogelijk bij de normering van advisering door algoritmes dan wel deskundigen.¹⁰³

5.2 Rationaliteit en het zorgvuldigheids- en gelijkheidsbeginsel

De inzet van deskundigenadvisering en beslissystemen om de rationaliteit van een beslissing te verhogen is vanuit de beginselen van behoorlijk bestuur wenselijk. Een bestuursorgaan moet immers zorgdragen voor een zorgvuldige vaststelling en kwalificatie van de feiten. Het zorgvuldigheids- en gelijkheidsbeginsel vergen dat die beoordeling zo geobjectiveerd mogelijk en consistent plaatsvindt. Zowel de deskundige als het beslissysteem kan overtuigend invulling geven aan die eisen.

100 ABRvS 27 juni 2018, ECLI:NL:RVS:2018:2086, r.o. 5.1, AB 2018/373, m.nt. A.C. Hendriks

101 Zie bijvoorbeeld iMMO, nieuwsbrief nr. 27, 10 december 2018, te raadplegen op: stichtingimmo.nl.

102 *PG Awb I*, p. 216-217; *Stcrt.* 2018, nr. 50 999 (Advies Raad van State betreffende digitalisering), p. 7.

103 Vgl. Wolswinkel bij een VAR-studiemiddag: M.J. Verzo, 'Bestuursrechtelijke uitdagingen in een data-gestuurde samenleving', *NTB* 2018/74, p. 479-480.

Over de band van objectiviteit kunnen beslissystemen op het eerste gezicht de voorkeur hebben boven individuele deskundigenadviezen. Breed geaccepteerd is dat algoritmes menselijke prestaties op sommige punten kunnen overtreffen.¹⁰⁴ Zo kunnen algoritmes een grote dataset over het algemeen beter analyseren op patronen en correlaties en beter leren van fouten. Van de uitkomsten van een algoritme kan handmatig of via *machine-learning* een analyse worden gemaakt van de ‘false positives’ en ‘false negatives’, waarna het systeem kan worden aangepast. Deskundige mensen daarentegen kunnen maar lastig leren van fouten.¹⁰⁵ ‘*Humans struggle to correct patterns of ineffective or biased decisions, often due to the lack of meaningful feedback, whereas algorithms are specifically designed to learn through error.*’¹⁰⁶ Vanuit dat gezichtspunt is het lastig te begrijpen waarom bij de WOZ-waardevaststelling gedurende de procedure naar een meer subjectieve vaststelling wordt overgestapt.

Ook op het gebied van consistentie lijkt het algoritmische beslissysteem aan de winnende hand. Het consistent toepassen van algemene (beslis)regels geeft op het niveau van het systeem gelijkheid en levert voor beleid belangrijke informatie op over de feitelijke uitvoering.¹⁰⁷ Dat benadrukte de minister destijds ook in de Tweede Kamer toen het CBBS ter discussie stond. Het systeem maakt meer inzichtelijk of de claimbeoordelingen per regio, kantoor of medewerker verschillen en genereert zo informatie aan de hand waarvan beleid kan worden bijgesteld.¹⁰⁸ Dit is een voordeel ten opzichte van bijvoorbeeld de planschadepraktijk waarin consistente advisering door gebrek aan informatie en de focus op (lokaal) maatwerk lastig is.¹⁰⁹

Toch ligt het genuanceerder. Het neutrale jargon van geautomatiseerde besluitvorming kan ook schijnobjectiviteit opleveren, omdat het verhult dat hierin allerlei onzekerheden en keuzes zijn verdisconteerd. Techniek is niet neutraal. Voor de ontwikkeling van het systeem moet de systeemontwikkelaar een doel formuleren, datasets selecteren, variabelen kiezen en getalsmatig duiden en een model bouwen, waarin wettelijke kaders en beleid worden omgezet in zwart-wit redeneringen.¹¹⁰ Bij al deze elementen kunnen bewust of onbewust onzuivere of onjuiste keuzes worden gemaakt, wat ten koste gaat van de accuraatheid en doelmatigheid van het beslissysteem. Ook kunnen datasets zelf gebiased zijn, waardoor de uitkomsten uit het systeem dat ook zijn en kunnen discrimineren.¹¹¹ In dat opzicht verschilt geautomatiseerde besluitvorming niet fundamenteel van deskundigenadvisering. In beide gevallen worden beleidsmatige keuzes gemaakt en moeten bijvoorbeeld mechanismen worden ingebouwd om fouten te kunnen corrigeren

104 Bijv. R.H. Sloan & R. Warner, 'When is an Algorithm Transparent? Predictive Analytics, Privacy and Public Policy', *IEEE Security & Privacy* 2018, afl. 3, p. 18-25.

105 Behalve als mensen in een voorspelbare omgeving kunnen oefenen en feedback ontvangen, volgens D. Kahneman, *Ons feilbare denken (Thinking, fast and slow)*, Amsterdam: Uitgeverij Business Contact 2015, p. 254-255.

106 Oswald 2018, p. 15.

107 *Sicr.* 2018, nr. 50 999 (Advies Raad van State betreffende digitalisering), p. 3.

108 *Kamerstukken II* 2005/06, 28 333, nr. 72.

109 Zie voor een zeldzaam geslaagd beroep op het gelijkheidsbeginsel: ABRvS 3 oktober 2018, ECLI:NL:RVS:2018:3185, *Gst.* 2019/21, m.nt. J.W. van Zundert.

110 Zie o.a. Diakopoulos 2016, p. 56-62.

111 Zie nader Max Vetzo, Janneke Gerards & Remco Nehmelman, *Algoritmes en grondrechten*, Den Haag: Bju 2018; F. Zuiderveen Borgesius, *Discrimination, artificial intelligence, and algorithmic decision-making* (rapport Raad van Europa), Straatsburg 2018, p. 10-12.

en neutraliteit (van betrokken softwareontwikkelaars, programmeurs, ambtenaren of deskundigen) te waarborgen, zoals bij deskundigenadvisering al gebeurt door normen over onpartijdigheid.

Dit leidt tot een tweede nuance, die ziet op de verschuiving van het niveau van de individuele beoordeling naar het systeemniveau.¹¹² Negatief is dat de uniformerende werking van systemen oneerlijk kan uitwerken in het individuele geval, waartegen een burger lastig op kan komen en waarbij de menselijke maat soms verdwijnt.¹¹³ Positief daarentegen is dat meer aandacht voor de waarborgen op systeemniveau een impuls kan geven aan de rechterlijke toetsing van individuele besluiten gebaseerd op deskundigenadvisering.¹¹⁴ Bestaan in een bepaalde discipline 'systeemnormen' om fouten in beoordelingen te detecteren, dan versterkt dat de mate van deugdelijkheid van de adviezen. Zijn er bijvoorbeeld (evidence-based) protocollen voor de beroepsgroep, opleidingseisen en een tuchtrechtregime, dan vergroot dat de kans dat 'false positives' en 'false negatives' in de beoordelingen worden gecorrigeerd.¹¹⁵ De rechter zou deze factoren kunnen verdisconteren in de intensiteit waarmee hij deskundigenoordelen toetst. Gedeeltelijk zie je dat bij de FMMU-casus, waarin de bestuursrechter de geprotocolleerde beoordeling toetst en overigens ook de tuchtrechter zich bezighoudt met beoordelingen van FMMU-personeel.¹¹⁶ Zijn er weinig systeemnormen en krijgt het deskundigenadvies in de besluitvorming veel gewicht, dan vraagt de selectie en werkwijze van de deskundige bijzondere rechterlijke aandacht.

Een belangrijke vervolgvraag is welke beoordelingen zich lenen voor algoritme-gedreven besluitvorming. Algoritmes in het publieke domein lijken thans vooral toegepast te worden om risicoanalyses uit te voeren in het toezichtstraject.¹¹⁷ Risicoanalyses hebben echter veel bredere potentie, ook in het kader van vergunningverlening. Zo kan voor een verlenging van een rijbewijs een risicoanalyse worden gebruikt om de rijvaardigheid van iemand boven een bepaalde leeftijd in te schatten op basis van rijgedrag in de afgelopen jaren.¹¹⁸ Bij een ontheffing van de Flora- en Faunawet zou een risicoanalyse het effect van de ontheffing op de instandhouding van de soort kunnen aangeven. De

112 *Stcrt.* 2018, nr. 50 999 (Advies Raad van State betreffende digitalisering), p. 5; B.M.A. van Eck, M.A.P. Bovens & S. Zouridis, 'Algoritmische rechtstoepassing in de democratische rechtsstaat', *NJB* 2018/2101, p. 3010.

113 *Stcrt.* 2018, nr. 50 999 (Advies Raad van State betreffende digitalisering), p. 5; A. Tollenaar, 'ICT en de uitdagingen voor de menselijke maat', in: A.T. Marseille & L. van der Velden (red.), *Verdient vertrouwen, vertrouwen verdient: Visies op geschilbeslechting door de overheid*, Den Haag: Sdu uitgevers 2014, p. 120-129.

114 Zie hierover ook Oswald 2018, p. 6.

115 Deskundigheid is namelijk ook iets collectiefs, gebaseerd op gedeelde kennis, waarden en methoden, zie M. Ambrus e.a., 'The role of experts in international and European Decision-making processes: setting the scene', in: M. Ambrus e.a. (eds.), *The Role of 'Experts' in International Decision-Making Processes: Advisors, Decision Makers or Irrelevant Actors?*, Cambridge University Press 2014, p. 12.

116 Zie bijvoorbeeld CTG 17 april 2018, ECLI:NL:TGZCTG:2018:107, waarin een FMMU-verpleegkundige een waarschuwing kreeg voor het onvoldoende uitvragen van psychische klachten.

117 S. Doove & D. Otten, *Verkenkend onderzoek naar het gebruik van algoritmen binnen overheidsorganisaties*, Den Haag: CBS 2018.

118 Liefst op basis van data afkomstig van de elektronische auto waarmee hij of zij de afgelopen jaren heeft gereden. Vgl. het voorbeeld van C. Coglianese & D. Lehr, 'Transparency and Algorithmic Governance', *Public Law and Legal Theory Research Paper Series, Research Paper No. # 18-38* (ssrn), te verschijnen in *Administrative Law Review*, p. 37.

mogelijkheden tot algoritmische ondersteuning lijken vrij onbeperkt, de crux zit vooral in elementen zoals het ontwerp van het systeem, de foutenmarge, de instructie aan de gebruikers van het systeem en de mate waarin zij uitkomsten nog handmatig kunnen bewerken. Toegepast op de casus van het CBBS was bijvoorbeeld cruciaal of de getalsmatige duiding van de normaalwaarden wetenschappelijk onderbouwd kon worden en welke gebruikersinstructie de verzekeringsarts en arbeidsdeskundige kregen over de noodzaak van invoer van bijzondere gegevens of handmatige bewerking van de resultaten. Bij het PAS blijkt onder het voorzorgsbeginsel de wetenschappelijke onderbouwing van onzekerheidsmarges van groot belang. De vraag in welke gevallen bestuursorganen gebruik mogen maken van algoritmische advisering betreft ook niet enkel een juridische vraag. Daaraan ligt tevens een politieke en bestuurlijke afweging ten grondslag, met name welke foutenmarges en risico's (zoals op verkeersongevallen en uitsterven van soorten) maatschappelijk aanvaardbaar zijn en hoe dit opweegt tegen waarden zoals autonomie en privacy.

5.3 Accountability en het motiverings- en equality of arms beginsel

Vanuit het bestuursrecht is het uitgangspunt bij zowel deskundigenadvisering als (deels) geautomatiseerde besluitvorming duidelijk: het bestuursorgaan is verantwoordelijk voor het genomen besluit en moet dit besluit motiveren om controle van overheidshandelen mogelijk te maken, ongeacht of het hulpmiddelen heeft ingeschakeld. De bestuursrechter moet vervolgens garanderen dat de burger die opkomt tegen het besluit gelijke proceskansen heeft.

Dit heldere juridische uitgangspunt toont zich voor besluiten op basis van algoritmes in eerste instantie weerbarstig: het ontwerp en de werking van een beslissysteem zijn vaak te complex om te begrijpen hoe tot een specifiek besluit is gekomen. Vaak is nog wel mogelijk om technische informatie te ontsluiten, zoals de broncode, de variabelen of gebruikte data.¹¹⁹ Deze 'technische transparantie' biedt alleen meestal niet de oplossing voor de motivering van een besluit, aangezien enerzijds de meeste ontvangers de technische deskundigheid missen om deze informatie te begrijpen en anderzijds volledige informatie risico's van manipulatie (*gaming*) kan meebrengen.¹²⁰ Meer technisch onderlegde auteurs voegen hieraan toe dat volledige (technische) transparantie vaak ook geen nuttige informatie oplevert over de oordeelsvorming achter de individuele beslissing die uit het systeem rolt.¹²¹ Maakt een systeem gebruik van *machine learning* dan is volledige technische transparantie vaak niet eens mogelijk, omdat het systeem zichzelf al

119 Coglianese & Lehr (nog te verschijnen), p. 27; Zie *Kamerstukken II*, 2018/19, 26 643, nr. 570, waarin de Minister voor rechtsbescherming de mogelijkheden tot technische transparantie duidt.

120 C. Castelluccia & D. Le Métayer, *Understanding algorithmic decision-making: Opportunities and challenges* (rapport STOA), Maart 2019, p. 77; D.R. Desai & J.A. Kroll, 'Trust but Verify: A guide to Algorithms and the law', *Harvard Journal of Law & Technology* 2018, afl. 1, p. 7-8; A.D. Selbst & S. Barocas, 'The Intuitive Appeal of Explainable Machines', *Fordham Law Review* 2018, p. 1085-1139, p. 1093.

121 Desai & Kroll 2018, p. 10; M. Ananny & K. Crawford, 'Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability', *New Media & Society* 2018, p. 973-989.

lerende aanpast.¹²² Zelfs de systeemontwikkelaar heeft dan geen (volledig) inzicht in de werking van deze systemen.¹²³ De grote nadruk die de Nederlandse rechtspraak legt op *equality of arms*, ook bij het adviserende algoritme, verbaast in dat licht.¹²⁴ Het is zeer de vraag of een belanghebbende in staat is dergelijke complexe materie, zelfs met behulp van deskundigen, te betwisten.

Om *accountability* mogelijk te maken en vertrouwen in algoritme-gedreven besluitvorming te stimuleren, worden vanuit zowel de juridische als technische hoek oplossingen voorgesteld.¹²⁵ Om zicht te krijgen op wat de overheid doet, is het vooral belangrijk dat deze uitlegt welke redenen een uitkomst rechtvaardigen en wat de logica is van het systeem.¹²⁶ Dit hangt bijvoorbeeld af van welke beleidsmatige doelen zijn gesteld, welke (feitelijke) gegevens zijn meegenomen, welke menselijke betrokkenheid er (mogelijk) is, welke foutenmarge het systeem kent en welke keuzes zijn gemaakt voor de weging van variabelen tot het uiteindelijke resultaat.¹²⁷ Zo heeft ook de Afdeling in het toetsingskader bij het PAS niet de nadruk gelegd op de technische kant van het algoritme; zij heeft zich vooral geconcentreerd op de onderbouwing van de beleidsmatige keuzes, de gebruikte aannames en gegevens. Ontvangers van een geautomatiseerd genomen besluit willen meestal ook niet de technische kant van een beslissysteem betwisten, maar de wijze waarop de regels in dit geval zijn toegepast of waarom dít de regels zijn. Het kan ook niet van belanghebbenden of de rechter worden verwacht de kwaliteit van het algoritme - zonder deskundige - te controleren. Wat wel zou kunnen is dat het bestuursorgaan informatie ontsluit over de wijze waarop het zelf de accurateheid van het algoritme heeft getest door bijvoorbeeld de test- en trainingsdata of betrouwbaarheidsanalyses te verstrekken.¹²⁸ Ook kan het nuttig zijn niet zozeer informatie over de werking van 'de black box' te geven, maar juist over bij de ontwikkeling betrokken actoren en de procedure die is gevolgd om het systeem aan te laten sluiten bij de context waarin het wordt gebruikt.¹²⁹ Ook meer technische oplossingen kunnen soelaas bieden, zoals software die in voor mensen begrijpelijke taal kan uitleggen hoe de analyse heeft plaatsgevonden.¹³⁰ Afhankelijk van de context zouden dergelijke methodes de controleerbaarheid van het algoritme kunnen bevorderen.

Ook de *accountability* van menselijke deskundigenadviezen verdient kritische reflectie. Waar de feitenkwalificatie een zekere prognose inhoudt (zoals waardedaling van de onroerende zaak of de resterende verdien capaciteit) worden inschattingen gemaakt die niet volledig te objectiveren zijn. Zeker als het aankomt op ervaring en intuïtie leent een

122 Coglianese & Lehr (nog te verschijnen), p. 13.

123 Selbst & Barocas 2018, p. 1094.

124 En niet op het motiveringsbeginsel, zoals ook Assink noemt in zijn noot bij ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259, *Gst.* 2017/170.

125 Zie over beide: Selbst & Barocas 2018, p. 7-12.

126 R. Binns, 'Algorithmic accountability and public reason', *Philosophy & Technology* 2018, p. 543-556; Selbst & Barocas 2018, p. 1125.

127 In de kamerbrief (*Kamerstukken II*, 2018/19, 26643, nr. 570) noemt de minister dit 'uitlegbaarheid'. Zie ook Diakopoulos 2016, p. 59-61.

128 Coglianese & Lehr (nog te verschijnen), p. 33-34; Desai & Kroll 2017, p. 64-65.

129 Annany en Crawford doen de suggestie om niet de werking van de systemen te bekijken, maar het netwerk waarin ze worden gebruikt, Annany & Crawford 2018, p. 974.

130 Coglianese & Lehr (nog te verschijnen), p. 41.

beoordeling zich evenmin als een complex algoritme voor een inzichtelijke ontleding in alle (technische) denkstappen. Dat lijkt de bestuursrechter van deskundigen ook niet te vragen, hoewel ook hier een beweging naar meer kritische toetsing zichtbaar is.¹³¹ In wetenschap en rechtspraak wordt gezocht naar mogelijkheden om deskundigenadviezen zinvol 'te controleren', bijvoorbeeld door transparantie te verwachten in de vorm van deskundigenregisters en 'disclosure statements' waarin de deskundige transparant moet zijn over opleiding, ervaring en (potentiële) belangenverstrengelingen.¹³² Voor *door de rechter* ingeschakelde deskundigen zijn ondertussen verschillende gedragscodes opgesteld met onder andere transparantieverplichtingen over de gebruikte methode waaraan deze deskundige moet voldoen, die wellicht ook bij deskundigenadvisering in de bestuurlijke fase kunnen worden toegepast.¹³³ Zodoende is ook hier nog werk aan de winkel voor het ontwikkelen van behulpzame mechanismen om *accountability* te versterken.¹³⁴

6 Conclusie

Software die het bestuursorgaan 'adviseert' om een bepaald besluit te nemen, zet het bestuursrecht aan opnieuw te doordenken hoe ver de rechter moet gaan in het openen van black box besluiten, aan de hand van welke normen en wat daarvan de ratio is. In deze bijdrage hebben we daartoe een nadere aanzet willen geven.

Het zorgvuldigheids- en motiveringsbeginsel en het beginsel van *equality of arms* geven de bestuursrechter instrumenten om de rechtmatigheid van algoritme-gedreven of op deskundigenadviezen gebaseerde besluitvorming te beoordelen. Voor beide hulpmiddelen geldt dat de bestuursrechter niet in staat is een inhoudelijk waardeoordeel te vellen over de juistheid van het gegeven advies: het is veelal te complex en valt buiten de deskundigheid van de rechter. De rechter zoekt dan al snel de waarborgen in meer procedurele eisen, zoals de eis dat belanghebbenden over de onderliggende gegevens moeten kunnen beschikken en hen de gelegenheid moet worden geboden om tegenbewijs te leveren. Die processuele waarborgen versterken vooral de inzichtelijkheid en controleerbaarheid van het advies en versterken *accountability*. Bij deze op zich positieve ontwikkeling passen twee kritisch kanttekeningen. Ten eerste moet de aandacht voor inzichtelijkheid en controleerbaarheid niet ten koste gaan van de rationaliteit. Een beoordeling op meer absolute waarden en heldere categorieën (zoals de 28 aspecten binnen het FIS, puntensystemen in subsidietenders of de eis dat het immo kenbaar moet maken op welk onderdeel van het asielrelaas de psychische beperking invloed heeft gehad) maken de waardering transparant, maar kunnen leiden tot schijnobjectiviteit, waarvoor de deskundige vanuit zijn vakgebied soms geen verantwoordelijkheid kan

131 Zie bijvoorbeeld Jaarverslag Raad van State 2017, Den Haag 2018, p. 60.

132 B.J. van Ettekoven, 'De deskundige deskundige over registers en de 'disclosure statement'', *O&A* 2016/53.

133 Zie de Gedragscode voor gerechtelijk deskundigen bij de Afdeling bestuursrechtspraak (*Stcr.* 2018, nr. 6861, 12 februari 2018) en de Gedragscode voor gerechtelijke deskundigen in civielrechtelijke en bestuursrechtelijke zaken. Zie N.H. van Amerongen & R. Kegge, 'De Gedragscode in het omgevingsrecht. Een oplossing voor gedragsproblemen bij de deskundige?', *TBR* 2019/30.

134 C. Holst & A. Molander, 'Public deliberation and the fact of expertise: making experts accountable', *Social Epistemology* 2017, afl. 3, p. 235.

nemen. Het kan dan realistischer zijn om te erkennen dat ervaring en intuïtie een rol spelen of handmatige aanpassing mogelijk moet zijn. Ten tweede vormen ‘gelijke processuele kansen’ al snel een onhaalbaar ideaal in het geval van complexe algoritmische analyses of specialistische adviezen, waarvoor de last veelal terechtkomt op de schouders van de burger. Voor een zinvolle discussie over de deugdelijkheid van het advies zijn ook andere waarborgen nodig.

Die andere waarborgen zouden wat ons betreft bijvoorbeeld gezocht kunnen worden in de organisatie van de advisering dan wel de ontwikkeling van het algoritme. De rechter duikt niet in de inhoudelijke wetenschappelijke of professionele discussie, of de mathematische analyse van het beslissysteem, maar beziet met welke waarborgen het systeem van oordelen is omgeven. De rechter zou bij deskundigenadvisering factoren kunnen meewegen als welke professionele standaarden en opleidings- en scholingseisen bestaan en in welke mate feedbacksystemen bestaan om foutieve oordelen te corrigeren, zoals tuchtrecht. Bij het beslissysteem zou de rechter kunnen meewegen wie welke keuzes heeft gemaakt en welke betrouwbaarheidsanalyses het bestuursorgaan heeft uitgevoerd om de deugdelijkheid van het algoritme te monitoren.

Uiteindelijk menen wij dat de toetsing van op deskundigenadvisering dan wel algoritmische analyse gebaseerde *black box* besluiten slechts gradueel verschilt en dat beide toetsingskaders elkaar zouden kunnen versterken. Zo kan van de toetsing van algoritmische besluitvorming worden geleerd hoe bepalend het ontwerp van het adviesstelsel is voor de uitkomst in de individuele zaak. De kritische benadering die beslissystemen ondertussen ontvangen past ook de organisatie van menselijke advisering. Van de toetsing van deskundigenadvisering kan het belang van de selectie van de deskundige en diens (on)partijdigheid worden meegenomen. Mens en machine (het systeem) zijn uiteindelijk sámen deskundig.